UNCLASSIFIED

AD 4 20 0 05

DEFENSE DOCUMENTATION CENTER

FOR

SCIENTIFIC AND TECHNICAL INFORMATION

CAMERON STATION, ALEXANDRIA, VIRGINIA

UNCLASSIFIED

CONVEX PROGRAMMING—
DUAL ALGORITHM

by

Pierre Huard
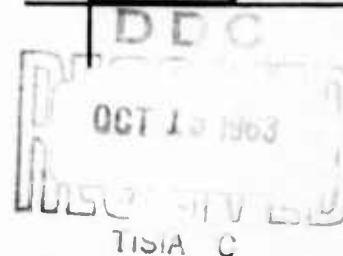
OPERATIONS RESEARCH CENTER

INSTITUTE OF ENGINEERING RESEARCH

UNIVERSITY OF CALIFORNIA - BERKELEY

CONVEX PROGRAMMING - DUAL ALGORITHM

by

Pierre Huard[*]
Operations Research Center
University of California, Berkeley

16 July 1963

ORC 63-21 (RR)

ABSTRACT

This work is the result of a study of G. B. Dantzig's [1] algorithm
for nonlinear programs where the objective as well as inequality constraints
are convex functions. This algorithm consists of finding an equivalent
"generalized linear program" where the standard procedure for determining
the candidate to enter the basis is replaced by an auxiliary nonlinear pro-
gram where the nonlinear constraints of the original problem have disappeared.
The algorithm described in this paper is very close to Dantzig's algorithm,
but the relations between the two algorithms are not yet well established.

I. Definitions-Properties

The initial program

PROBLEM 1:
$$\begin{cases} \text{Find } x \text{ such that } \phi_o(x) \text{ is maximum} \\ \text{subject to } \phi_i(x) \geq 0 \qquad i = 1,2,\ldots,m \ . \end{cases}$$

where ϕ_o and ϕ_i are continuously differentiable concave functions in x .

We will call $\phi(x)$ or ϕ the vector whose elements are $\phi_i(x)$.

We assume that the optimal value of $\phi_o(x)$ is finite.

Let $\pi = (\pi_1, \pi_2, \ldots, \pi_m)$ be a row vector.

Consider the following program:

PROBLEM 3:
$$\begin{cases} \text{Find } x \text{ such that } \psi(x,\pi) \text{ is maximum} \\ \text{where } \psi(x,\pi) = \pi\phi(x) + \phi_o(x) \text{ and } \pi \text{ is given } (\pi \geq 0) \ . \end{cases}$$

Let $z(\pi)$ be an optimal solution of Problem 3.

Assumption 1: $z(\pi)$ is continuous in π and finite for all $\pi \geq 0$.

Let

(1) $\Theta(\pi) = \pi\phi(z) + \phi_o(z)$.

THEOREM 1: The gradient of $\Theta(\pi)$ in the π-space is equal to $\phi(z)$.

PROOF: At the optimum, $z(\pi)$, of Problem 3 we have

(2) $\dfrac{\partial\psi(x,\pi)}{\partial x} = 0$ (Maximum with respect to x)

so, at $x = z(\pi)$

-1-

$$(3) \qquad \frac{d\Theta(\pi)}{d\pi} = \frac{\partial \psi(x,\pi)}{\partial x} \frac{dz}{d\pi} + \frac{\partial \psi(x,\pi)}{\partial \pi} = \phi(z) .$$

THEOREM 2: $\qquad \Theta(\pi) = \psi(x,\pi)$ is a convex function in π .

PROOF: Let λ' and λ'' be two scalars such that $\lambda' + \lambda'' = 1$, and λ' , $\lambda'' \geq 0$. Let π' , $\pi'' \geq 0$ be two values of π ; and z' , z'' the corresponding optimal values for Problem 3. Then,

$$(4) \qquad \Theta'' - \Theta' = \pi''\phi'' + \phi''_0 - (\pi'\phi + \phi'_0)$$

$$\geq \pi''\phi' + \phi'_0 - (\pi'\phi' + \phi'_0)$$

$$\text{because } \Theta(\pi'') = \underset{x}{\text{Max}} \; \{\psi(x,\pi'')\}$$

$$= (\pi'' - \pi') \; \phi'$$

or $\qquad \Theta'' - \Theta' \geq (\pi'' - \pi') \left[\dfrac{d\Theta}{d\pi}\right]_{\pi=\pi'} \qquad$ by Theorem 1
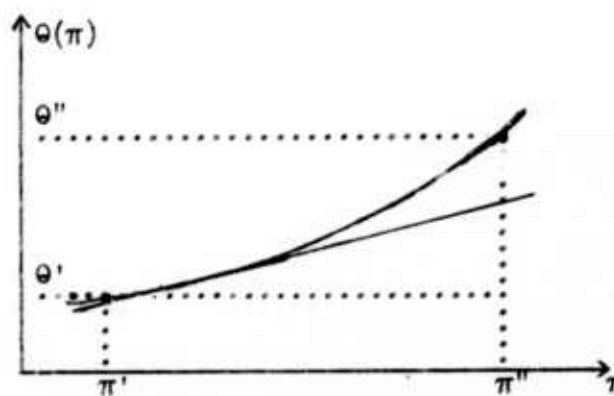

Figure 1

i.e., the variation of $\Theta(\pi)$ is always greater than or equal to that tangent plane, which characterizes a convex function.

Let the following program be:

PROBLEM 2:   [Find $\pi \geq 0$ such that $\Theta(\pi)$ is minimum

Assumption 2:   There exists a finite optimal solution $\hat{\pi}$ to Problem 2.

THEOREM 3:   The optimal value of the objective function of Problem 2 is equal to $\phi_0(\hat{x})$ the optimal value of the objective function of Problem 1.

PROOF: A necessary condition of optimality for Problem 2 is

$$(5) \qquad \pi \frac{d\Theta}{d\pi} = 0 \qquad \text{(Kuhn-Tucker condition)}$$

It follows that

$$(6) \qquad \hat{\pi}\phi(\hat{\pi}) = 0 ,$$

and

$$(7) \qquad \Theta(\hat{\pi}) = \phi_0(\hat{x}) .$$

II. The Algorithm

We replace the initial Problem 1 by Problem 2 of which the optimal solution yields the optimal solution for Problem 1.

Problem 2 is the minimization of a convex function, $\Theta(\pi)$, where the variable π is subjected to being nonnegative. The expression of the function $\Theta(\pi)$ is not given, but one can compute its value for all π , using the auxiliary program Problem 3. This auxiliary program gives us also the value, $\phi(z)$ of the gradient of $\Theta(\pi)$ at any given π .

We can solve Problem 2 as follows:

-3-

(1) Start with $\pi = \pi^k \geq 0$.

Compute $\Theta(^k\pi)$ and $\phi[z(\pi^k)]$ using Problem 3.

Determine the vector $g(\pi^k)$ defined by

(8)
$$\begin{cases} g_i = \phi_i[z(\pi^k)] & \text{if } \phi_i \leq 0 \underline{\text{ and/or}} \text{ if } \pi_i > 0 \ . \\[2em] g_i = 0 & \text{otherwise (i.e., if } \phi_i > 0 \underline{\text{ and }} \pi_i = 0) \ . \end{cases}$$

In order to simplify the notation g will be considered in some cases to be a row vector and in some other cases to be a column vector.

(2) On the half-line $\pi = \pi^k + \lambda g(\pi^k)$, $\lambda \geq 0$, determine a new point $\pi'^k \geq 0$ at a finite distance corresponding to $\lambda = \lambda'_k \geq 0$ and satisfying the condition

(9)
$$0 < \lambda'_k \leq \underset{i}{\text{Min}} \left\{ \frac{-\pi_i^k}{g_i} \ \middle| \ g_i < 0 \right\}$$

(3) Compute $\Theta(\pi'^k)$ using Problem 3

If $\Theta(\pi'^k) < \Theta(\pi^k)$, go to 1 , replacing π^k by $\pi^{k+1} = \pi'^k$.

If $\Theta(\pi'^k) \geq \Theta(\pi^k)$ and if π^k is not the optimal solution, we have necessarily

(10)
$$(\pi'^k - \pi^k) \quad \cdot \quad \phi(\pi^k) < 0$$

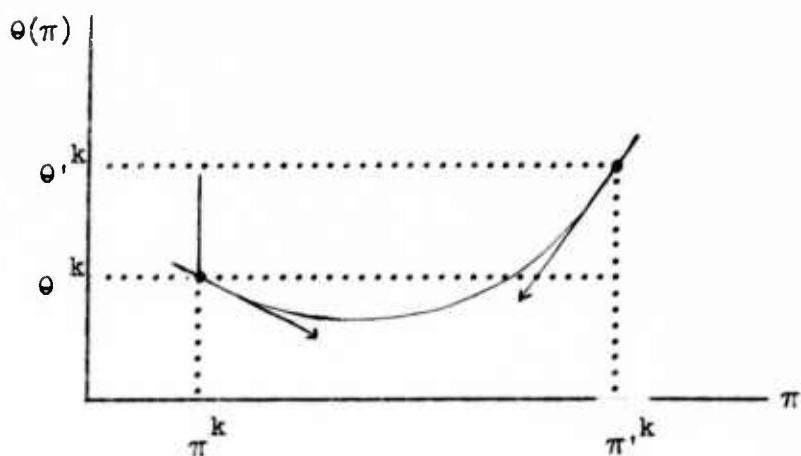$$(\pi'^k - \pi'^k) \quad \cdot \quad \phi(\pi'^k) > 0 \ .$$

Figure 2

Determine then a new point

(11) $\pi''^k = \pi^k + (1/2)\lambda'_k g(\pi^k)$

i.e., the middle point of (π^k , π'^k) or another intermediate point, taking into account the respective slopes at π^k and π'^k.

Then repeat ③ , with π''^k instead of π'^k, and so on until we obtain (in a finite number of steps) a point π^{k+1} such that

(12) $\Theta(\pi^{k+1}) < \Theta(\pi^k)$.

REMARK: To prove the convergence, we have to perform the subdivision of λ_k once more, which yields a point π^{*k+1} , and work with the better of π^{*k+1} and π^{k+1}. In order to simplify the notations, we will call the selected point π^{k+1}, and λ_k the value of λ for that point.

④ Go to ① with π^{k+1} instead of π^k .

-5-

III. Convergence

We can show easily (see ③ of the algorithm), that the variation of $\Theta(\pi)$, at each step k , is strictly negative, except if $g(\pi^k) = 0$ which implies that $\emptyset[z(\pi^k)] \geq 0$ and $\pi^k \cdot \emptyset[z(\pi^k)] = 0$. In these conditions, we achieved the optimum of Problem 2.

Since the minimum of $\Theta(\pi)$, $\pi \geq 0$ is equal to $\emptyset_o(\mathfrak{X})$, optimum of Problem 1 which is assumed to be finite, the sequence of points π^k gives a sequence of convergent values $\Theta(\pi^k)$ because we have a lower bound for these values. Moreover, we will make the following assumption:

Assumption 3: The second partial derivatives of $\Theta(\pi)$ are bounded above in all directions by a scalar $D > 0$.

One can show that:

$$(13) \qquad \Theta(\pi^k) - \Theta(\pi^{k+1}) \geq (1/2) \lambda_k \left| g^k \right|^2 \text{ if } \lambda_k \leq 1/D ,$$

$$(14) \qquad \Theta(\pi^k) - \Theta(\pi^{k+1}) \geq 1/4 \left| \frac{g^k}{D} \right|^2 \text{ if } \lambda_k \geq 1/D$$

where $\left| g \right|$ = modulus of g , letting $g^k = g(\pi^k)$.

Since λ_k is either determined by (g) or equal to a fraction of the first value λ_k' that we tried at the beginning of each step which can be taken equal to a constant h , for all k .

We obtain for all λ_k such that $0 \leq \lambda_k \leq h$.

$$(15) \qquad \Theta(\pi^k) - \Theta(\pi^{k+1}) \geq K \cdot \lambda_k \left| g^k \right|^2 ,$$

with $K = \text{Min} \{1/2 , 1/4hD\}$

from which we deduce that the series $\displaystyle\sum_k \lambda_k \left| g^k \right|^2$ converges and that

(16) $\qquad \lambda_k \, |g^k|^2 \to 0 \qquad$ as $\quad k \to \infty$

For each step k, by the convexity of $\Theta(\pi)$ we have

(17) $\qquad 0 \le \Theta(\pi^k) - \Theta(\hat\pi) \le (\pi^k - \pi) \cdot \phi^k$

where $\hat\pi$ is an optimal solution of Problem 2.

By definition of g^k:

(18) $\qquad (\pi^k - \hat\pi) \, \phi^k = - (\pi^k - \hat\pi) \, g^k + (\pi^k_N - \hat\pi_N) \cdot \phi^k_N$

where ϕ_N is a vector containing only the components of ϕ such that $\phi_i > 0$, $\pi_i = 0$. Then

(19) $\qquad (\pi^k_N - \hat\pi_N) \cdot \phi^k_N = -\hat\pi_N \cdot \phi^k_N \le 0$,

and finally

(20) $\qquad 0 \le \Theta(\pi^k) - \Theta(\hat\pi) \le -(\pi^k - \hat\pi) \, g^k$,

and in particular

(21) $\qquad (\pi^k - \hat\pi) \, \lambda_k \cdot g^k \le 0$

which means that the vector $(\pi^{k+1} - \pi^k)$ has an acute angle with $(\hat\pi - \pi^k)$

It follows

(22) $\qquad |\hat\pi - \pi^{k+1}|^2 \le |\hat\pi - \pi^k|^2 + |\pi^{k+1} - \pi^k|^2 =$

$\qquad\qquad |\hat\pi - \pi^k|^2 + \lambda_k^2 \, |g^k|^2$.

Summing with respect to k :

(23) $\qquad |\hat\pi - \pi^k|^2 \le |\pi - \pi^0|^2 + \sum_{k'=0}^{k} \lambda_{k'}^2 \, |g^{k'}|^2$

Since $\displaystyle\sum_{k'=0}^{k} \lambda_{k'} \, |g^{k'}|^2$ converges and $\lambda_{k'} \leq h$, the series

$$\sum_{k'=0}^{k}{}' \lambda_{k'} \, |g^{k'}|^2$$

converges when $k \to \infty$, from which we obtain an upper bound for the distance

$$|\pi^k - \hat{\pi}| \, .$$

We have to consider two cases:

CASE 1: $g^k \to 0$, at least for a finite or infinite subsequence of k . Then the corresponding subsequence

(24)
$$g^k(\pi^k - \hat{\pi}) \to 0 \quad \text{and} \quad \Theta(\pi^k \to \Theta(\hat{\pi})$$
For these conditions, the algorithm converges to an optimal solution.

CASE 2: g^k does not converge to zero for any subsequence of k . This means g^k is different from zero for all k , and moreover $|g^k|$ has a lower bound > 0 . For these conditions, the convergence of the series $\displaystyle\sum{}' \lambda_k \, |g^k|^2$ implies the convergence of the series $\displaystyle\sum \lambda_k \, |g^k|$ and $\displaystyle\sum \lambda_k$ and therefore implies the convergence of the point π^k which was defined by

(25)
$$\pi^k = \sum_{k'=0}^{k} \lambda_{k'} \, g^{k'}$$

Let $\hat{\pi}$ be the limit of π^k , and let $\tilde{g}$ be the limiting value of g^k at that point. (This limit exists by Assumption 1.)

The convergence of the series $\displaystyle\sum \lambda^k$ implies that $\lambda_k \to 0$ and by Assumption 3, λ_k is determined by the relation (9) for K sufficiently large. There exists at least one index i and two subsequences K' and K"

-8-

corresponding to the steps k such that,

$$(26) \quad k' \in K'' \begin{bmatrix} g_1^{k'} < 0 & ; & \phi_1^{k'} > 0 \\[2ex] \pi_1^{k'} > 0 & ; & \\[2ex] \lambda_{k'} = \dfrac{\pi_1}{-g_1} \end{bmatrix}$$

and

$$(27) \quad k'' \in K'' \begin{bmatrix} g_1^{k''} > 0 & ; & \phi_1^{k''} < 0 \\[2ex] \pi_1^{k''} = 0 \end{bmatrix}$$

So $\phi_1^{k'}$ and $\phi_1^{k''}$ go to zero, and we have by the fact that the second order derivatives of ϕ have an upper bound, D ,

$$(28) \qquad -g_1^{k'} \leq D \; . \; \pi_1^{k'}$$

If j is the index of a nonzero component of $\tilde{g}$, we have for k' sufficiently large,

$$(29) \qquad \lambda_{k'} \; |\, g_j^{k'} \,| = - \frac{\pi_1^{k'}}{g_1^{k'}} \, |\, g_j^{k'} \,| \geq \frac{1}{D} \, |\, g_j^{k'} \,| \cong \frac{1}{D} \, |\, \tilde{g}_j \,| > 0$$

We have here a contradiction with the fact that π^k converges to $\tilde{\pi}$ which implies that $\lambda_{k'} \; g_j^{k'} \to 0$. Thus the components of $\tilde{g}$ must be zero, which contradicts the assumption of a possible second case.

IV. REMARKS

1) Nature of the Algorithm: Our initial program is

Problem 1
$$\begin{bmatrix} \text{Find } x \text{ such that } \phi_o(x) \text{ is maximum} \\ \text{subject to } \phi_1(x) \geq 0 \end{bmatrix}$$

- 9 -

Let us consider the dual [2] of Problem 1:

$$(D) \quad \begin{cases} \text{Find } \pi \geq 0 \quad \text{such that } \psi(x,\pi) \text{ is minimum} \\ \text{subject to} \quad \dfrac{\partial}{\partial x_j} \psi(x,\pi) = 0 \qquad j = 1,2,\ldots,n \ . \end{cases}$$

We see that solving the auxiliary program Problem 3 for a given nonnegative π , is equivalent to finding a feasible solution (x,π) for the dual program (D) . Problem 2, which minimizes $\Theta(\pi)$ solves (D) by a discrete gradient method.

2) On Assumption 1: Assumption 1 seems at first very restrictive, nevertheless, if the optimum of Problem 1 is finite, one can always remove the constraints $\phi_i(x)$ and the functional $\phi_0(x)$ far enough from the origin in order to have $z(\pi)$ at a finite distance. For instance, this modification can consist in joining the surfaces $y_i = \phi_i(x)$ and $y_0 = \phi_0(x)$ to the vertical cylinder $|x| = \text{constant}$ if the constant is chosen large enough. This would reduce the search for an optimum of Problem 3 to the interior of a sphere, $(x) = \text{constant}$. One could imagine other compact spaces.

The continuity of $z(\pi)$ for π can be achieved, e.g., by assuming the optimum of Problem 3 is always unique. This is the case if $\psi(x,\pi)$ is strictly concave in x , i.e., if one at least of the functions $\phi_i(x)$ or $\phi_0(x)$ is strictly concave.

REFERENCES

1. Dantzig, G. B., _Linear Programming and Extensions_, Chapter 24, Princeton University Press, Princeton, New Jersey, 1963.

2. Huard, P., "Dual Programs," _Proceedings, Fifth Symposium on Mathematical Programming_, Chicago, Illinois, June 1962.

Professor A. Charnes
The Technological Institute
Northwestern University
Evanston, Illinois

Professor R. Radner
Department of Economics
University of California
Berkeley 4, California

Professor Murray Rosenblatt
Department of Mathematics
Brown University
Providence 12, Rhode Island

Mr. J. R. Simpson
Bureau of Supplies and Accounts
Navy Department (Code W31)
Washington 25, D. C.

Professor A. W. Tucker
Department of Mathematics
Princeton University
Princeton, New Jersey

Professor J. Wolfowitz
Department of Mathematics
Lincoln Hall, Cornell University
Ithaca 1, New York

C. O., ONR Branch Office
346 Broadway
New York 13, New York
    ATTN: J. Laderman

Professor Oskar Morgenstern
Economics Research Project
Princeton University
92 A Nassau Street
Princeton, New Jersey